CJFS 3710 Social Statistics OLS Regression Spring 2016



N=1,500

**OLS = Ordinary Least Squares** or ***Linear*** regression for short—it assumes that the DV is normally distributed as well as the IV, that the IV are not overly related to one another

**Dependent Variable (DV)** is “Should marijuana be legalized” (1=Yes, 0=No)

**Independent Variables (IV)** are:

 Age: # of years old the respondent is

 Educ: highest # of years the respondent has completed (2 years of college = 14)

 Sex : 1=Male, 2= Female

 Partners: # of sex partners the respondent had last year

**Standardized Regression Coefficients** allow you to compare within the equation (which is most predictive or powerful)

**Unstandardized Regression Coefficients** tell you for every 1 unit increase in the respective IV, how much change is there in the DV while taking into account all other IV

**Adjusted R Square** tells you the proportion of variance explained in the DV by all the IV in the regression equation

**Sig** = P Value or the proportion due to chance---the smaller Sig is, the less the var’s effect is due to something the researcher has no knowledge about—random error or chance

Sample Size needed: for each IV, you need 20 casesCJFS 3710 Social Statistics Logistic Regression Spring 2016





N=1,500

**Logistic Regression** or ***Nonlinear*** regression for short—it assumes that the DV is not normally distributed, is dichotomous, and that some of the IV can be skewed or categorical (nominal)

**DV and IV** are the same as in the OLS example

**B = Unstandardized Regression Coefficients** tell you for every 1 unit increase in the respective IV, how much change is there in the DV while taking into account all other IV

**Nagelkerke's R-Square** varies from 0 to 1 and is logistic’s version of R square in OLS

**Exp (B) = Odds Ratio**: when the 95% confidence interval (CI) around the odds ratio includes the value of 1.0 that indicates that a change in value of the IV is not associated in change in the odds of the DV assuming a given value (therefore, that IV is not considered a useful predictor)

**Sig** = P Value or the proportion due to chance (from the Wald statistic)

\*Note for a discussion of Logistic Regression see: <http://www.uwm.edu/~edari/methstat/logistic.htm#regress>